# SI Appendix

## SI Results and Discussion

***Isotope-labeled cross-linker enabled confident cross-link identification***—Cross-link data analysis is notoriously difficult owing to the large number of potential cross-linked peptides that might have formed and that must be included in the database search. This number rises roughly with the square of the number of proteins in the database. Distinguishing real cross-linked peptides from false positives thus becomes a key challenge. Our use of a 1:1 mixture of labeled and unlabeled cross-linker (1-3), based on a strategy described in (1), addressed this challenge by providing an easily-observed "fingerprint" in the liquid chromatogram, MS spectra, and MS/MS spectra indicating a real cross-link (see Figures 2,3, and S3). In the chromatogram, the light and heavy forms eluted off the LC column within ~10 s of each other; in the MS spectra, the light and heavy forms occurred as characteristic doublets with a 12 Da peak shift; and in the MS/MS spectra, the light and heavy forms were identical, except that product ions containing the cross-linker molecule occurred with a +12 Da shift for the heavy form. This shift in the MS/MS spectra also helped determine the amino acid sequence of product ions, by revealing whether or not a particular product ion contained the cross-linker molecule or not. All cross-linked peptides presented in this study displayed all of the characteristic fingerprints, and nearly all major product ions matched ones predicted for that peptide, resulting in highly confident cross-link identification.

The characteristic doublet observed in MS spectra of real cross-links addressed another key challenge in cross-link identification, namely that there tend to be many more non-linked peptides than cross-linked peptides present in a typical protein digest after cross-linking. With the typical "highest-abundance ion" criterion for selecting ions for fragmentation by MS/MS (a prerequisite for peptide sequence determination), many more non-linked peptides are selected than cross-linked peptides. This is a diversion of precious instrument time towards sequencing uninformative peptides. Enrichment of cross-linked peptides before LC-MS/MS helps, and chromatographic (4-6) or affinity (7-10) enrichment has been used successfully for this purpose. In this study, we performed effective enrichment of cross-linked peptides *during* MS using the "mass tags" feature of the instrument to screen each mass spectrum for isotopic doublets on-the-fly, and select only those doublets for fragmentation, as described by the Borchers group (1). This strategy permitted the instrument to ignore most uninformative non-linked peptides. Using this approach, we identified more PSII cross-links than we and others have found using the same cross-linker without this feature enabled (11).

***Evaluation of cross-link data quality—***As mentioned in the "Results," we measured the distance between the $C_\alpha$'s of linked residues in the PSII crystal structure (12) to assess data quality, using a 30 Å upper threshold for cross-links to be considered consistent with the crystal structure (Fig. S1, Table S2). This value was obtained as follows: with a lysine side chain distance of 6.5 Å and cross-linker arm length of 11.4 Å, the C-α's of linked residues should be within 24.4 Å of each other in the crystal structure. An additional allowance of several angstroms is typically made to account for protein flexibility, and a 30 Å threshold is used commonly (13-17). A recent survey study (18) confirmed this value to be appropriate.

Of the thirteen measurable cross-link distances, eleven (85%) were less than 30 Å (Fig. S1, Table S2), which matches well with the 89% value obtained in the recent large-scale study (18) of cross-links found in the *XLdb* (19). The remaining two cross-links had distances of 34 and 38 Å. In both of these cases, at least one of the two linked residues lies on a flexible loop of the protein, which could facilitate cross-linking over the slightly larger distance between the sites.

**SI Materials and Methods**

**Generation of the Psb28-His construct—**To generate the Psb28-His construct for expression in *E. coli*, the Psb28 gene was amplified from *Synechocystis* 6803 genomic DNA using the following primers: forward- 5′-GACATATGGCTGAAATTCAATTTTCCAAGG-3′; reverse- 5′-TCGGATCCTTAATGGTGATGGTGATGGTGATGGTGTTCAGATTTGGAAAAACCTAAG CCATTTTCTGCGCCG-3′. The forward primer contains an *Nde*I restriction site and the reverse primer contains a *Bam*HI restriction site. In the reverse primer, sequence encoding an 8x His-tag was inserted before the stop codon. The PCR fragment was cloned into the pET21a vector, and then transformed into the BL21(DE3) strain of *E. coli* for expression of Psb28-His.

**Purification of Psb28-His from *E. coli*—**The BL21(DE3) strain of *E. coli* containing the pET21a-Psb28-His vector was grown at 37 °C with constant shaking in 200 mL of LB medium containing 50 μg/mL ampicillin. When the culture reached $OD_{600}$ = 0.6, expression of Psb28-His was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The culture was harvested 3 h after induction by centrifugation at 3000 g for 15 min. Cells were resuspended by vortexing in 10 mL lysis buffer (20 mM Tris, pH 7.9, 100 mM NaCl, 10 mM imidazole), with 1 μg/mL DNase (Sigma, St. Louis, MO), 100 μL protease inhibitor cocktail (Sigma), and 1 mg lysozyme. After allowing cells to rock gently on ice for 15 min, they were lysed by probe sonication on ice for 3 min at 50% duty cycle, then centrifuged at 31,000 *g* for 15 min at 4 °C. The supernatant ("lysate") was applied onto a 1 mL pre-packed HisTrap FF column (GE Healthcare, Little Chalfont, Buckinghamshire, UK), pre-equilibrated with 10 mL lysis buffer, by dropwise manual injection. After collecting the flow-through, the column was washed with 10 mL lysis buffer. Elution was by manual injection of 10 mL elution buffer (lysis buffer containing 200 mM imidazole). Ten 1-mL fractions were collected, and analyzed by SDS-PAGE for Psb28 content. The most concentrated fraction contained 13 mg/mL Psb28 and was used for the subsequent experiments. A portion of this fraction was submitted to Cocalico Biologicals (Reamstown, PA) for generation of anti-Psb28 antisera in rabbit.

**Cyanobacterial culture and PSII purification—**Generation of the Δ*psb*O-His47 (20), Δ*ctp*A-His47 (21), Δ*ctp*A:Δ*psb27*-His47 (22), and His27 (21) strains has been reported previously. The HT3 (His47) strain was a kind gift from Dr. Terry Bricker (Louisiana State University, Baton Rouge, LA) (23). Cyanobacterial strains were grown in BG11 medium at 30 °C under 30 μmol photons $m^{-2} \cdot s^{-1}$. The growth media were supplemented with 10 μg/mL spectinomycin and 5 μg/mL kanamycin (Δ*psb*O-His47); 5 mM glucose, 10 μM 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), 5 μg/mL kanamycin, and 3 μg/mL erythromycin (Δ*ctp*A-His47); 5 mM glucose, 10 μM DCMU, 10 μg/mL chloramphenicol, and 5 μg/mL kanamycin (Δ*ctp*A:Δ*psb27*-His47); and 5 μg/ml gentamicin (His27). Purified PSII samples were stored in 25% glycerol, 10 mM $MgCl_2$, 5 mM $CaCl_2$, 50 mM MES buffer (pH 6.0) (RB buffer). For glycerol gradient ultracentrifugation, the

linear gradient was made by using stock solutions of 5% and 30% weight/vol glycerol in RB buffer. PSII samples in RB buffer were diluted to 5% glycerol before ultracentrifugation, and 50 or 100 μg Chl *a* containing samples were loaded in each tube. PSII monomer and dimer represented by their bands were recovered after ultracentrifugation and were concentrated using Vivaspin 500 centrifugal concentrators (100,000 molecular weight cutoff) (Vivaproducts, Littleton, MA). Protein gel electrophoresis and immunoblotting were performed as described previously (24, 25). Immunoblot quantification analysis was performed with the ImageQuant TL software package (GE Healthcare, Pittsburgh, PA).

**Cross-linking and proteolytic digestion—**The $BS^3$ cross-linker was dissolved in a stock solution of RB buffer. This solution was added to aliquots of PSII samples containing 1 or 2 μg Chl *a* (in RB buffer after ultracentrifugation and concentration) at a cross-linker:PSII molar ratio of 50:1, 100:1, and 300:1, and incubated in the dark at room temperature for 50 min. The reaction was quenched by addition of a solution containing 1 M Tris (pH 7.5) to 50 mM final concentration. Salt and detergent were removed from the cross-linked samples by precipitation using the 2D-Clean-Up kit (GE Healthcare). Cross-linked protein pellets were resuspended in 20 μL 8 M urea, 50 mM ammonium bicarbonate. A two-step digestion with lysyl endopeptidase (LysC) and trypsin was used to increase final yield of tryptic peptides (4). Lysyl endopeptidase (MS-grade, Wako Chemicals USA, Richmond, VA) was added to the protein sample at a 1:50 LysC:protein ratio (weight:weight) and the samples were incubated at 37 °C for 2 h. After 2 h, samples were diluted 1:8 in 50 mM ammonium bicarbonate to a final urea concentration of 1 M. Trypsin (Sigma, St. Louis, MO) was added to the samples at a 1:25 trypsin:protein ratio (weight:weight) and the samples were incubated at 37 °C overnight. After digestion, samples were acidified to a final concentration of 1% formic acid and analyzed by LC-MS/MS.

**LC-MS/MS—**Aliquots (5-μL) of peptide samples were loaded onto an Ultimate 3000 Nano LC system (Thermo Scientific Dionex, Sunnyvale, CA) attached in-line to a Q Exactive Plus mass spectrometer (Thermo Fisher, Waltham, MA). Peptide samples initially flowed through a guard column (Acclaim PepMap100, 100 μm × 2 cm, C18, 5 μm, 100 Å; Thermo Scientific Dionex) in Solvent A (water with 0.1% formic acid) and were separated on a $C_{18}$ reversed-phase column (Magic, 0.075 mm ×150 mm, 5 $\mu$m, 120 Å, Michrom Bioresources, Inc., Auburn, CA) packed in house, at 4.5 μL/min. Peptides were eluted using a linear 90-min gradient from 5-95% solvent B (80% acetonitrile, 20% water, 0.1% formic acid), followed by a 10-min hold at 95% solvent B. Eluted samples were flowed directly into the mass spectrometer via a PicoView Nanospray Source (PV550, New Objective, Inc., Woburn, MA) with a spray voltage of 1.8 kV. The instrument was operated in positive-ion mode with a scan range from *m/z* 380-1500. Full mass spectra were acquired at 70,000 resolving power for ions of *m/z* 200, with automatic gain control set at $3 \times 10^6$ ions and a maximal injection time of 200 ms. Data-dependent product-ion spectra were acquired at 17,500 mass resolving power for ions of *m/z* 200, with automatic gain control set at $1 \times 10^5$ ions and a maximal injection time of 100 ms. The top 15 precursor ions were fragmented by HCD with an isolation window of 3.0 *m/z* and normalized collision energy of 30%. For samples run with the "mass tags" feature enabled (1), only precursor ion pairs with a mass-to-charge shift selection criteria of ±6.03762, ±4.02508, ±3.01881, ±2.41505, ±2.01254, or ±1.72503 Da, corresponding to peptide charge states of 2-7, respectively, were selected for fragmentation. Up to the ten top precursor ions matching these criteria were fragmented by HCD with an isolation window of 3.0 *m/z* and normalized collision energy of 25%. Non-cross-linked samples were analyzed excluding

charge states other than 2-7, and each cross-linked sample was analyzed in duplicate or triplicate, in runs that excluded charge states other than either 2-7, 3-7, and 4-7. Data presented in Fig. S3 were acquired on an LTQ-Orbitrap XL (Thermo Fisher, Waltham, MA), as described previously (20), with the "mass tags" feature enabled as described above.

**Cross-link data analysis—**The raw LC-MS/MS files were searched for proteins in the *Synechocystis* proteome using PEAKS (ver. 7.0, Bioinformatics Solutions, Inc., Waterloo, ON, Canada). Peptides were identified with a 0.1% false discovery rate. Mono-linked peptides were identified by PEAKS (as well as by ICC-CLASS described below), with the light and heavy forms of the cross-linker as user-defined modifications (+156.0079 and +168.1540 Da, respectively). The identified proteins served as the database for cross-link detection using the ICC-CLASS software suite (1, 26). Raw LC-MS/MS files were converted to MGF format using Proteome Discoverer (Thermo Scientific) and to Xtract text files using the Xcalibur software File Converter Tool (Thermo Scientific). Isotopic doublets were detected in the MS data with the ICC-CLASS programs, and cross-link candidates were identified with DXMSMSMatchESI (1, 26). Search settings were as follows: Crosslinker- DSS; $M_{ip}$-137.06025; Mclrest1- 0 Da; Mclrest2- 0 Da; Digest sites- KR; Including CL site- No; Missed Digests- up to 2; CL sites- K; Dead-end peptides only- No; Intra-peptide only- No; Systematic error- 0 ppm; Mass tolerance- ± 10 ppm; Precursor tolerance- ± 10 ppm; Retention time tolerance- ± 2 s; Fragments tolerance- ± 10-500 ppm. After identification of cross-links using these search conditions, the search tolerance was expanded to identify cross-links that may have been missed initially by including CL site- Yes; Missed digests- up to 4. Product-ion spectra of candidate cross-links were inspected manually to verify the identification, using the predicted fragment ion masses in DXMSMSMatchESI or the PeptidesCL program (www.creativemolecules.com). Additional cross-link data analysis was performed using Protein Prospector (27, 28).

**Protein-protein docking—**The DOT program (29) uses convolution methods to perform a systematic rigid-body translational and rotational search. The van der Waals and electrostatic energies of both the molecules were mapped onto grids to predict the interactions between two macromolecules. One of the two molecules was rotated and translated around the other to predict the energetically favorable complexes. The Psb28 molecule was selected as the moving molecule and used a cubic grid of 192 Å on a side with 1 Å spacing between points for the translational search. A set of 54,000 rotational orientations, with a reduced over-sampling at the corners of the cubical grid, provided a resolution of 6° for the rotational search. Combined translational and rotational search resulted in over 382 billion configurations ($192^3$ * 54,000) of the PSII and Psb28 molecules. The top 4,000 docked conformations based on calculated interaction energy were saved and analyzed. These complexes formed four distinct clusters of which three were located in the membrane-spanning regions or on the luminal surface of PSII, making these clusters physiologically irrelevant; they were not considered further. The cluster localized at the cytosolic surface of PSII (Cluster 1) was composed of 1,282 Psb28 conformations. The top 100 conformations from Cluster 1 were investigated further as described in the Results section. Analysis of the docked conformations was conducted in VMD (30).

**SI References**


1. Petrotchenko EV, Makepeace KAT, Serpa JJ, Borchers CH (2014) Analysis of protein structure by cross-linking combined with mass spectrometry. *Methods Mol Biol* 1156:447-463.

2. Müller DR, et al. (2001) Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal Chem* 73(9):1927–1934.

3. Pearson KM, Pannell LK, Fales HM (2002) Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun Mass Spectrom* 16(3):149–159.

4. Leitner A, et al. (2012) Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol Cell Proteomics* 11(3):M111.014126. DOI: 10.1074/mcp.M111.014126.

5. Buncherd H, Roseboom W, de Koning LJ, de Koster CG, de Jong L (2014) A gas phase cleavage reaction of cross-linked peptides for protein complex topology studies by peptide fragment fingerprinting from large sequence database. *J Proteomics* 108:65-77

6. Fritzsche R, Ihling CH, Götze M, Sinz A (2012) Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun Mass Sp* 26(6):653-658.

7. Paramelle D, Miralles G, Subra G, Martinez J (2013) Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* 13(3-4):438-456.

8. Petrotchenko EV, Serpa JJ, Borchers CH (2011) An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol Cell Proteomics* 10(2): 10.1074/mcp.M110.001420.

9. Kang S, et al. (2009) Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry. *Rapid Commun Mass Spectrom* 23(11):1719-1726.

10. Chu F, Mahrus S, Craik CS, Burlingame AL (2006) Isotope-coded and affinity-tagged cross-linking (ICATXL): An efficient strategy to probe protein interaction surfaces. *J Am Chem Soc* 128(32):10362-10363.

11. Cormann KU, Möller M, Nowaczyk MM (2016) Critical assessment of protein cross-linking and molecular docking: An updated model for the interaction between Photosystem II and Psb27. *Front Plant Sci* 7:157.

12. Hellmich J, et al. (2014) Native-like Photosystem II superstructure at 2.44 Å resolution through detergent extraction from the protein crystal. *Structure* 22(11):1607-1615.

13. Shi Y, et al. (2015) A strategy for dissecting the architectures of native macromolecular assemblies. *Nat Methods* 12(12):1135-1138.

14. Zelter A, et al. (2015) The molecular architecture of the Dam1 kinetochore complex is defined by cross-linking based structural modelling. *Nat Commun* 6:8673.

15. Fischer L, Chen ZA, Rappsilber J (2013) Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J Proteomics* 88:120-128.

16. Herzog F, et al. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337(6100):1348-1352.

17. Walzthoeni T, et al. (2011) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* 9(9):901-903.

18. Merkley ED, et al. (2014) Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci* 23(6):747-759.

19. Kahraman A, et al. (2013) Cross-link guided molecular modeling with ROSETTA. *PLoS ONE* 8(9):e73411.

20. Liu H, et al. (2014) MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. *Proc Natl Acad Sci USA* 111(12):4638-4643.

21. Liu H, Roose JL, Cameron JC, Pakrasi HB (2011) A genetically tagged Psb27 protein allows purification of two consecutive photosystem II (PSII) assembly intermediates in Synechocystis 6803, a cyanobacterium. *J Biol Chem* 286(28):24865-24871.

22. Liu H, Huang RY-C, Chen J, Gross ML, Pakrasi HB (2011) Psb27, a transiently associated protein, binds to the chlorophyll binding protein CP43 in photosystem II assembly intermediates. *Proc Natl Acad Sci USA* 108(45):18536-18541.

23. Bricker TM, Morvant J, Masri N, Sutton HM, Frankel LK (1998) Isolation of a highly active Photosystem II preparation from *Synechocystis* 6803 using a histidine-tagged mutant of CP 47. *Biochim Biophys Acta* 1409(1):50–57.

24. Kashino Y, et al. (2002) Low-molecular-mass polypeptide components of a Photosystem II preparation from the thermophilic cyanobacterium *Thermosynechococcus vulcanus*. *Plant Cell Physiol* 43(11):1366–1373.

25. Kashino Y, Koike H, Satoh K (2001) An improved sodium dodecyl sulfate-polyacrylamide gel electrophoresis system for the analysis of membrane protein complexes. *Electrophoresis* 22(6):1004–1007.

26. Petrotchenko EV, Borchers CH (2010) ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics* 11:64.

27. Trnka MJ, Baker PR, Robinson PJJ, Burlingame AL, Chalkley, RJ (2014) Matching cross-linked peptide spectra: Only as good as the worse identification. *Mol Cell Proteomics* 13(2):420-434.

28. Chalkley RJ, Trnka MJ, Michael N, Baker, PR. (2014) Identifying cross-linked peptides using Protein Prospector. 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, June 15.

29. Roberts VA, Thompson EE, Pique ME, Perez MS, Ten Eyck LF (2013) DOT2: Macromolecular docking with improved biophysical models. *J Comput Chem* 34(20):1743-1758.

30. Humphrey W, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics. *J Molec Graphics* 14(1):33-38.

31. Chang L, et al. (2015) Structural organization of an intact phycobilisome and its association with photosystem II. *Cell Res* 25(6):726-737.

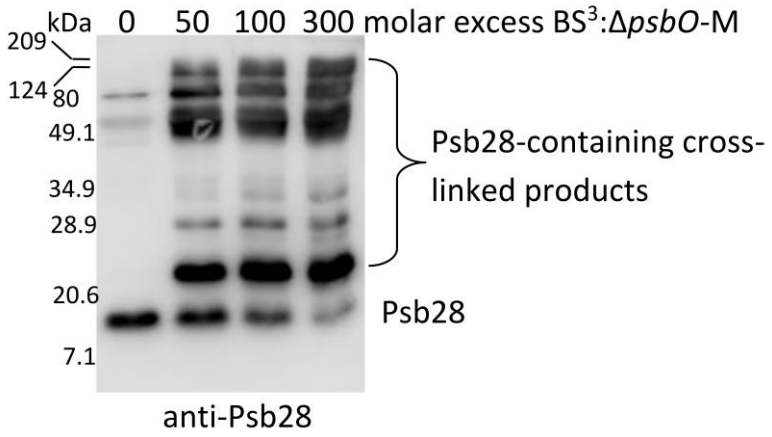



**Fig. S1** Immunoblot comparing Δ*psbO*-PSII before and after cross-linking with 50, 100, and 300 molar excess $BS^3$:PSII.

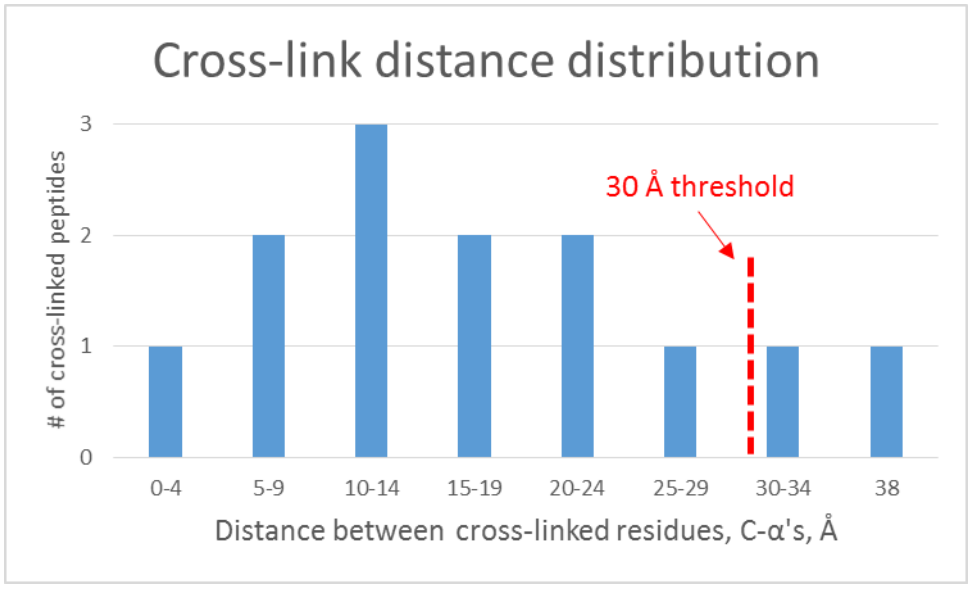



**Fig. S2** Distribution of linear (Euclidean) distances between C-α's of cross-linked residues in Δ*psbO*-His47-PSII.

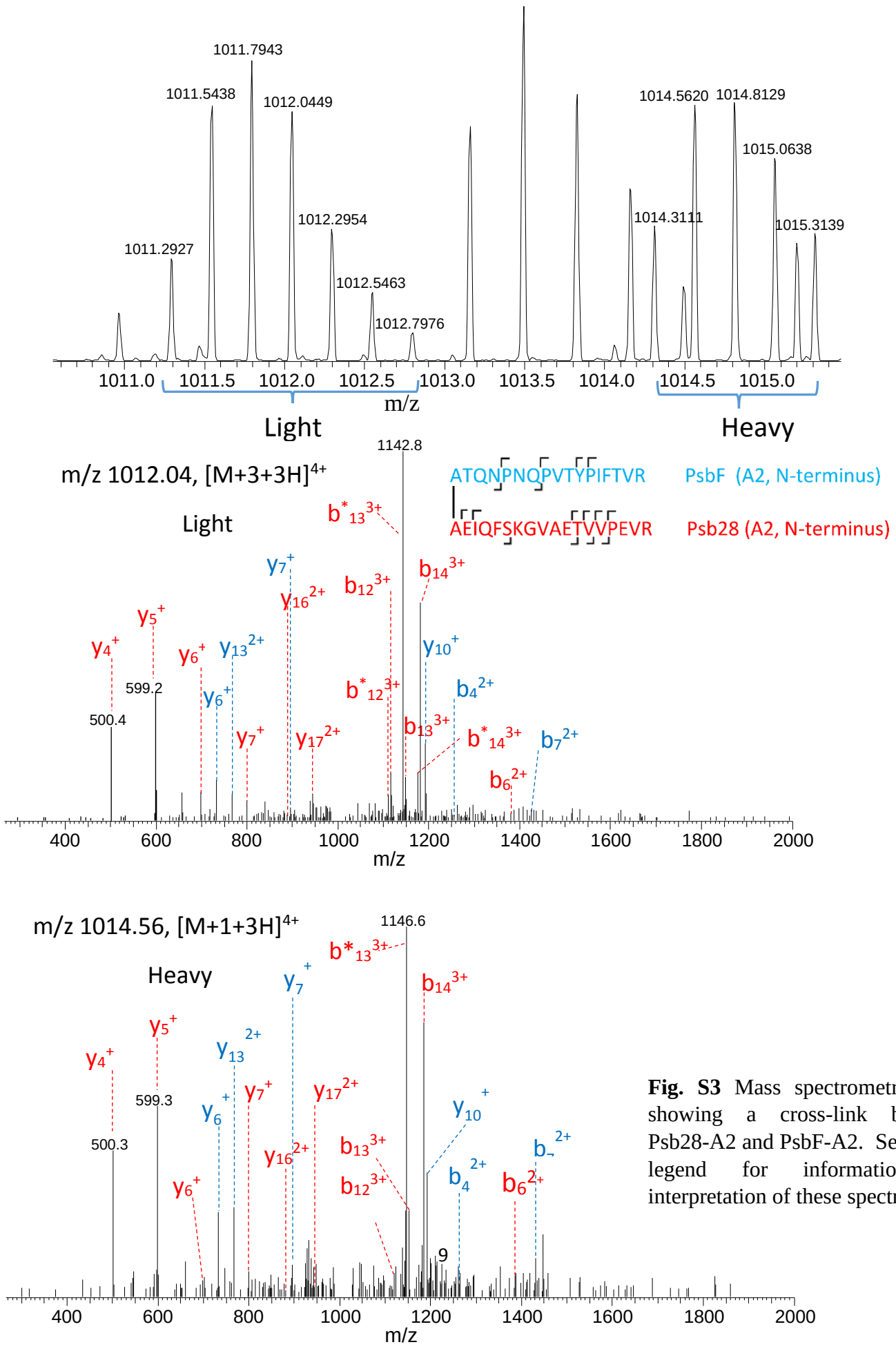



**Fig. S3** Mass spectrometric data showing a cross-link between Psb28-A2 and PsbF-A2. See Fig. 2 legend for information on interpretation of these spectra.

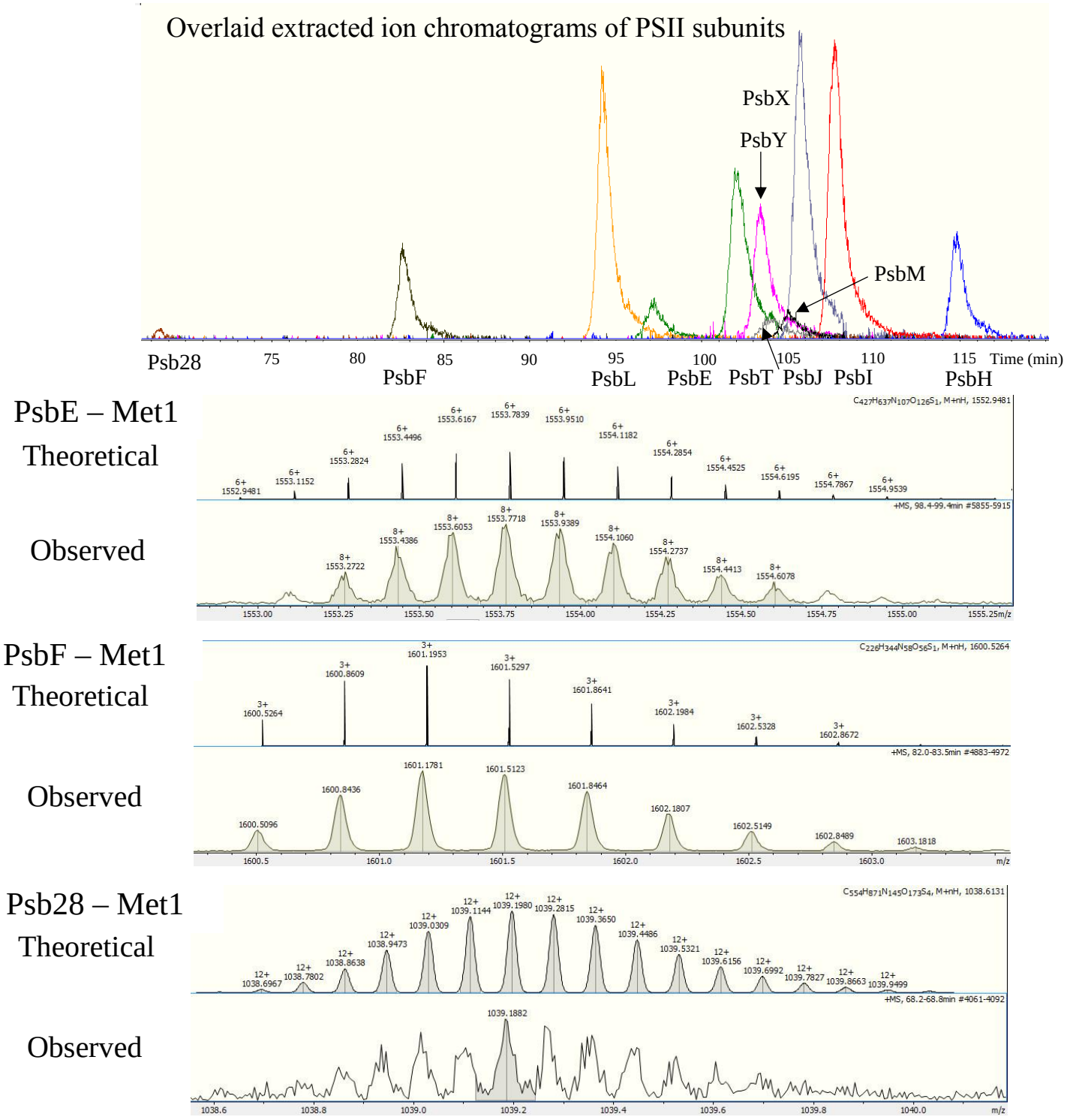
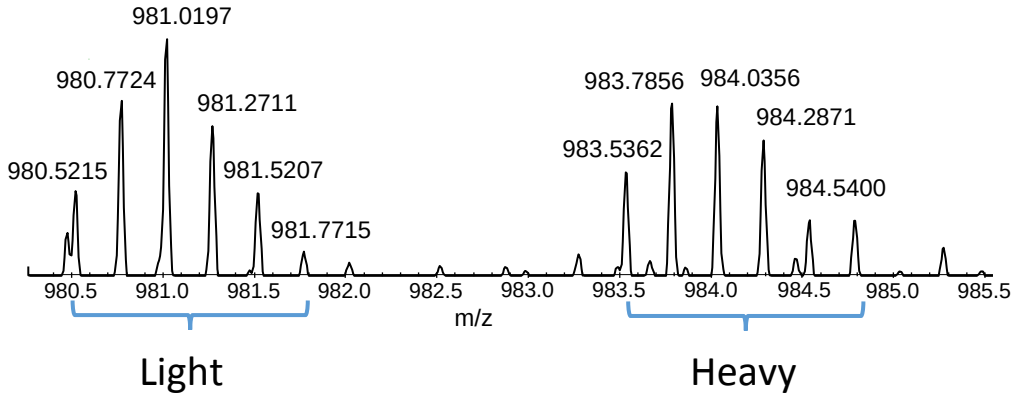



**Fig. S4** Intact-mass spectra of PsbE, PsbF, and Psb28 show that the N-terminal methionine has been cleaved for all three proteins, rendering PsbE-S2, PsbF-A2, and Psb28-A2 as the N-terminal residues of the mature proteins. Theoretical mass spectra of each protein are shown for comparison. Mass accuracy of 8, 11, and 10 ppm (within 0.1 Da) was achieved for PsbE, and PsbF, and Psb28, respectively.

Δ*psb*O-His47-PSII




His47-PSII

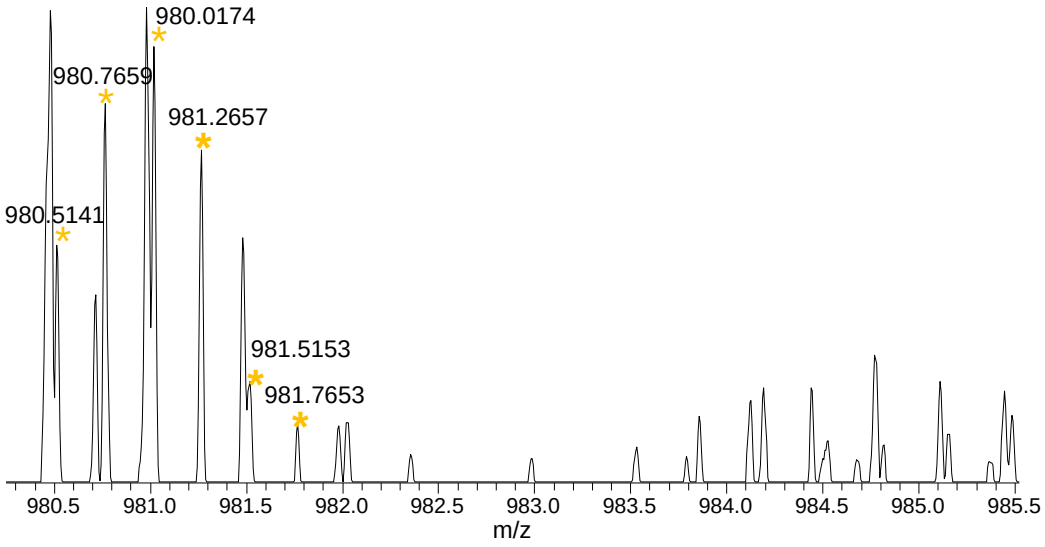



**Fig. S5** PsbE-Psb28 cross-link identified initially in Δ*psb*O-His47-PSII (upper spectrum) was also detected in His47-PSII (wild-type with a polyhistidine tag on the CP47 subunit) (lower spectrum, 3 ppm MS-1 mass accuracy). This demonstrates that the association we describe between Psb28 and PsbE is not an artifact of the absence of PsbO in Δ*psb*O-His47-PSII. The lower spectrum reflects a technical replicate experiment in which non-isotope-encoded $BS^3$ was used.

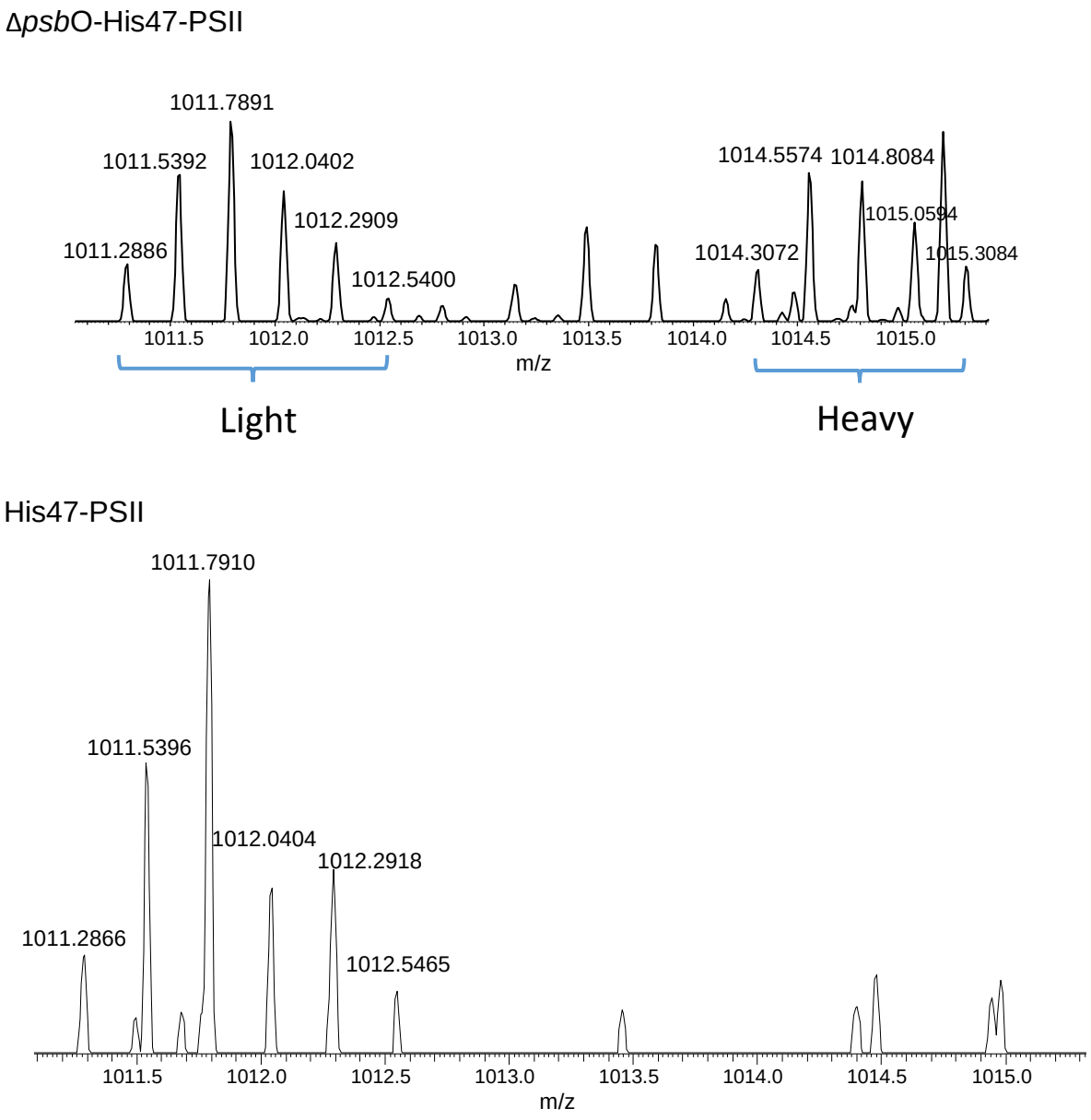
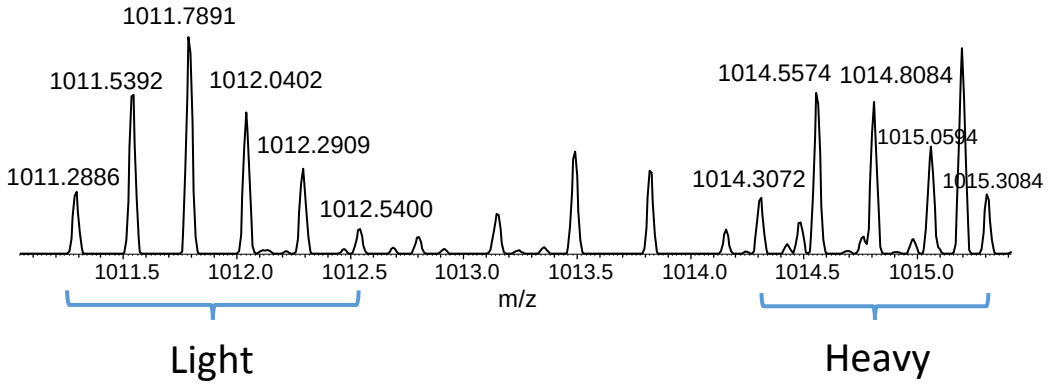



**Fig. S6** PsbF-Psb28 cross-link identified initially in Δ*psb*O-His47-PSII (upper spectrum) was also detected in His47-PSII (wild-type with a polyhistidine tag on the CP47 subunit) (lower spectrum, 1 ppm MS-1 mass accuracy). This demonstrates that the association we describe between Psb28 and PsbF is not an artifact of the absence of PsbO in Δ*psb*O-His47-PSII. The lower spectrum reflects a technical replicate experiment in which non-isotope-encoded $BS^3$ was used (the same technical LC-MS/MS experiment from which the lower spectrum in Fig. S5 was observed).

Δ*psb*O-His47-PSII




His47-PSII

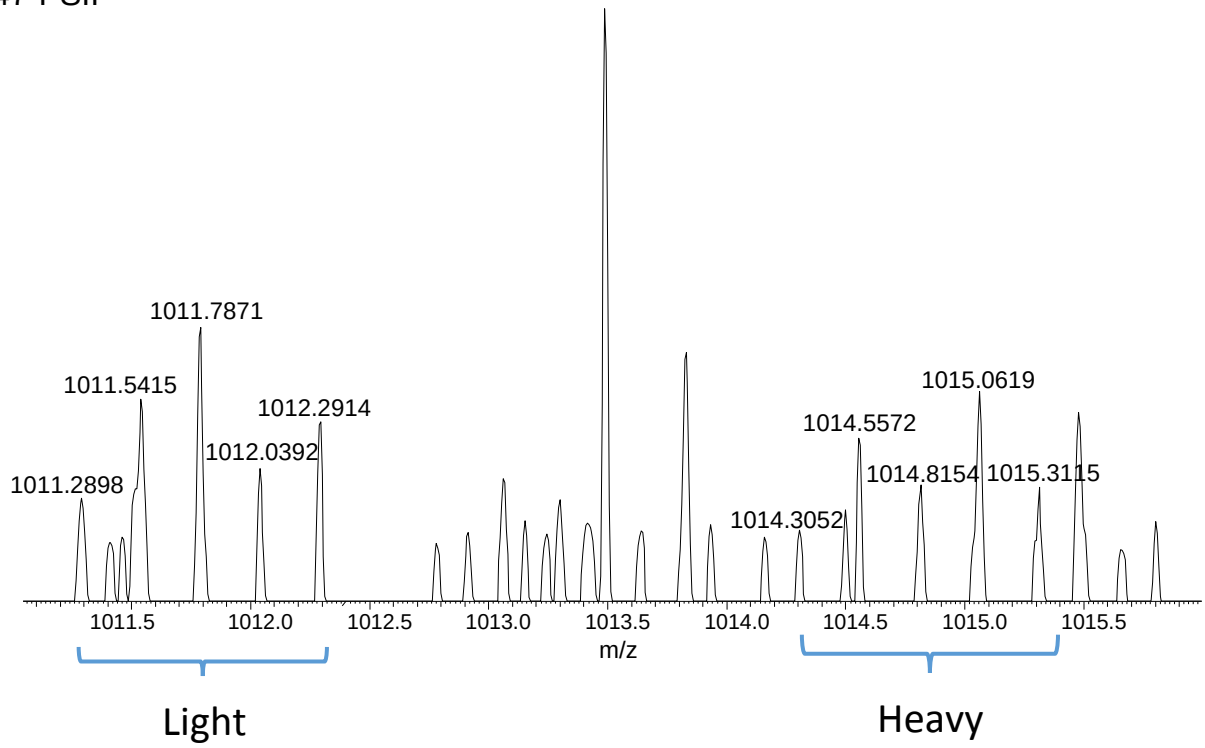



**Fig. S7** Same result as Fig. S6 (with 2 ppm MS-1 mass accuracy of the cross-link detected in His47-PSII), but in this figure the lower spectrum reflects a technical replicate experiment in which isotope-encoded $BS^3$ was used.

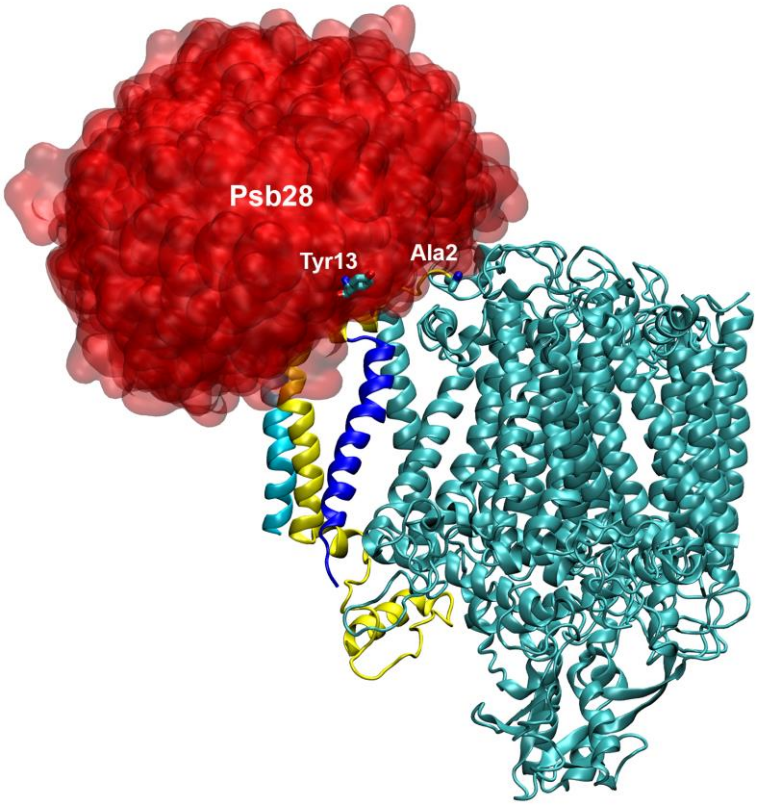



**Fig. S8** The top 100 Cluster-1 conformations show Psb28 binding to the RC47 complex on the cytosolic surface above PsbE (yellow) and PsbF (blue). PsbE-A2 and PsbF-Y13 are shown. PsbF-Y13 is actually obscured behind part of the Psb28 cloud, but its position has been brought to the fore in this image for clarity.

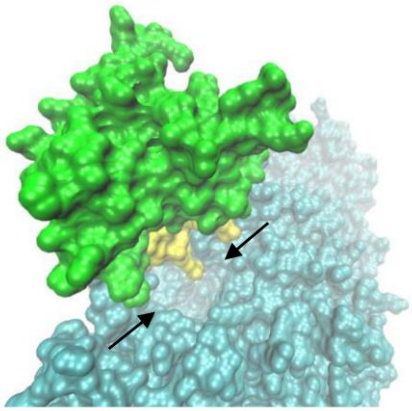


**Fig. S9** In our model Psb28 binds just beside the cavity on the cytosolic surface of PSII (shown by arrows) that is the site of phycobilisome attachment to PSII (31). Green, Psb28; yellow, PsbE; Cyan, RC47 proteins.

**Table S1** –Mono-linked peptides identified in this study.

| | [M+H]+ | Charge | Protein | Peptide sequence | Linked residue |
|---|---|---|---|---|---|
| **1** | 3872.7867 | +4 | PsbA | ETTEVESQNYGY**K**FGQEEETYNIVAAHGYFG R | K238 |
| **2** | 947.5002 | +2 | PsbB | **G**LPWYR | G2 (N-term) |
| **3** | 919.5618 | +2 | PsbB | LY**K**ALR | K227 |
| **4** | 2045.9609 | +3 | PsbB | YQWD**K**GYFQEEIQR | K277 |
| **5** | 2628.2678 | +3 | PsbB | TGAMNSGDGIAQEWIGHPIF**K**DK | K347 |
| **6** | 3752.8297 | +3 | PsbB | SES**K**FSVEQTGVTVSFYGGALDGQTFSNPSDV KK | K389 |
| **7** | 3320.6418 | +3 | PsbB | FSVEQTGVTVSFYGGALDGQTFSNPSDV**K**K | K418 |
| **8** | 2636.2192 | +3 | PsbB | **K**AQLGEGFDFDTETFNSDGVFR | K423 |
| **9** | 3137.5638 | +3 | PsbB | DVFAGVDPGLEEQVEFGVFA**K**VGDLSTR | K497 |
| **10** | 2174.1668 | +3 | PsbC | LGANIASAQGPTGLG**K**YLMR | K338 |
| **11** | 2302.2587 | +3 | PsbC | GPWLEPLRGPNGLDLD**K**LR | K378 |
| **12** | 1361.7065 | +2 | PsbC | AAAAGFE**K**GIDR | K456 |
| **13** | 1979.0105 | +3 | PsbE | **S**GTTGERPFSDIVTSIR | S2 (N-term) |
| **14** | 2102.0938 | +3 | PsbF | **A**TQNPNQPVTYPIFTVR | A2 (N-term) |
| **15** | 1730.9315 | +3 | PsbH | LGDILRPLNSEYG**K** | K20 |
| **16** | 3440.7029 | +3 | Psb27 | **K**GDAGGL**K**SFTTMQTALNSLAGYYTSYGAR | K56, K63 |
| **17** | 980.6027 | +2 | Psb27 | PIPE**K**LK | K90 |
| **18** | 1264.7966 | +3 | Psb27 | PIPEKL**K**KR | K92 |
| **19** | 978.5148 | +2 | Psb28 | **A**EIQFSK | A2 (N-term) |
| **20** | 2115.1369 | +3 | Psb28 | AEIQFS**K**GVAETVVPEVR | K8 |
| **21** | 2272.1636 | +3 | Psb28 | NGQSGMA**K**FYFLEPTILAK | K32 |
| **22** | 2968.5532 | +3 | Psb28 | G**K**FINGRPTAIEATVILNSQPEWDR | K69 |

**Table S2** –Cross-linked peptides identified in this study.

| | [M+H]+ | Charge | Protein 1 | Protein 2 | Cross-linked peptide sequence | Linked residues | Distance (Å) |
|---|---|---|---|---|---|---|---|
| **1** | 4645.1652 | +4 | PsbA | PsbB | ETTEVESQNYGY**K**FGQEEETYNIVAAHGYF GR—**G**LPWYR | K238-G2 (N-term) | 17 |
| **2** | 4391.9917 | +4 | PsbA | PsbI | ETTEVESQNYGY**K**FGQEEETYNIVAAHGYF GR—**K**DFE | K238-K35 | 34 |
| **3** | 3882.0473 | +4 | PsbB | PsbB | LY**K**ALR— DVFAGVDPGLEEQVEFGVFA**K**VGDLSTR | K227-K497 | 20 |
| **4** | 4338.1007 | +4 | PsbB | PsbB | SES**K**FSVEQTGVTVSFYGGALDGQTFSNPS DVKK—D**K**EGR | K349-K389 | 14 |
| **5** | 3677.8117 | +4 | PsbB | - | FSVEQTGVTVSFYGGALDGQTFSNPSDV**KK** FAR | K418-K419 | 4 |
| **6** | 3138.5333 | +4 | PsbB | PsbB | **K**FAR—**K**AQLGEGFDFDTETFNSDGVFR | K419-K423 | 7 |
| **7** | 2475.3975 | +5 | PsbB | PsbH | LY**K**ALR—LGDILRPLNSEYG**K** | K227-K20 | 12 |
| **8** | 1321.7312 | +3 | PsbB | PsbL | LY**K**ALR—**M**DR | K227-M1 (N-term) | 38 |
| **9** | 1349.6652 | +2 | PsbB | PsbL | **G**LPWYR—**M**DR | G2 (N-term)-M1 (N-term) | 24 |
| **10** | 1880.9339 | +3 | PsbC | PsbI | AAAAGFE**K**GIDR—**K**DFE | K456-K35 | 17 |
| **11** | 3906.0209 | +4 | PsbE | PsbF | **S**GTTGERPFSDIVTSIR— **A**TQNPNQPVTYPIFTVR | S2 (N-term)-A2 (N-term) | - |
| **12** | 3919.0502 | +4 | PsbE | Psb28 | **S**GTTGERPFSDIVTSIR— AEIQFS**K**GVAETVVPEVR | S2 (N-term)-K8 | - |
| **13** | 4042.1490 | +4 | PsbF | Psb28 | **A**TQNPNQPVTYPIFTVR— **A**EIQFSKGVAETVVPEVR | A2 (N-term)-A2 (N-term) | - |
| **14** | 4042.1354 | +4 | PsbF | Psb28 | **A**TQNPNQPVTYPIFTVR— AEIQFS**K**GVAETVVPEVR | A2 (N-term)-K8 | - |
| **15** | 3267.6082 | +3 | Psb27 | - | **K**GDAGGL**K**SFTTMQTALNSLAGYYTSYGA R | K56, K63 | 14 |
| **16** | 962.59099 | +2 | Psb27 | - | PIPE**K**L**K** | K90, K92 | 6 |
| **17** | 2058.0421 | +3 | Psb28 | Psb28 | LSKS**K**—YGAENGLGFS**K**SE | K24-K110 | 28 |
| **18** | 4273.2531 | +4 | Psb32 | Psb32 | TGVYDLPILSPGS**K**TFLVDQAEAISLANENR —LNSDL**K**K | K60-K83 | - |

The Psb28-PsbE and Psb28-PsbF cross-links, as well as the cross-linked residues for each entry, are shown in red. All distances are linear (Euclidean) distances between the C-α's of the two linked residues, as measured in PDB 4PJ0 (PSII), 2KND (Psb27) and 2KVO (Psb28) with adjustments for the following entries:

2, 10- PDB 3WU2 was used because PsbI-K35-C-α is not resolved in PDB 4PJ0.

3-PsbB-K497 corresponds to K498 in 4PJ0, so that residue was used for distance measurement.

5,6-PsbB-K419 corresponds to S419 in PDB 4PJ0, and that residue was used for distance measurement.

8,9- PsbL-M1 was not resolved in PDB 4PJ0, but was resolved in PDB 3WU2, so 3WU2 was used instead for these distance measurements.

10-PsbC-K456 corresponds to K457 in PDB 4PJ0, and that residue was used for distance measurement.

**Table S3** Energetics obtained from the Psb28-RC47 docking calculations for the top 100 Cluster-1 conformations. All energy values are given in kcal/mol.

| Conformation # | Interaction Energy | E_elec[1] | E_vdw[2] | Conformation # | Interaction Energy | E_elec[1] | E_vdw[2] |
|---|---|---|---|---|---|---|---|
| **1** | -56.11 | -52.41 | -3.70 | **51** | -45.41 | -39.51 | -5.9 |
| **2** | -55.04 | -44.14 | -10.90 | **52** | -45.31 | -43.31 | -2 |
| **3** | -53.91 | -44.91 | -9.00 | **53** | -45.25 | -39.85 | -5.4 |
| **4** | -52.37 | -45.17 | -7.20 | **54** | -45.23 | -33.43 | -11.8 |
| **5** | -52.27 | -51.77 | -0.50 | **55** | -45.16 | -40.46 | -4.7 |
| **6** | -52.20 | -40.20 | -12.00 | **56** | -45.09 | -36.79 | -8.3 |
| **7** | -51.64 | -43.04 | -8.60 | **57** | -45.04 | -37.74 | -7.3 |
| **8** | -50.92 | -45.72 | -5.20 | **58** | -45.01 | -40.21 | -4.8 |
| **9** | -50.86 | -41.46 | -9.40 | **59** | -45.01 | -37.51 | -7.5 |
| **10** | -50.68 | -43.38 | -7.30 | **60** | -44.99 | -37.19 | -7.8 |
| **11** | -50.52 | -47.22 | -3.30 | **61** | -44.94 | -37.94 | -7 |
| **12** | -49.57 | -38.77 | -10.80 | **62** | -44.87 | -39.17 | -5.7 |
| **13** | -49.50 | -42.40 | -7.10 | **63** | -44.84 | -41.74 | -3.1 |
| **14** | -49.25 | -42.05 | -7.20 | **64** | -44.77 | -37.77 | -7 |
| **15** | -49.05 | -40.55 | -8.50 | **65** | -44.67 | -38.97 | -5.7 |
| **16** | -48.68 | -45.58 | -3.10 | **66** | -44.63 | -35.73 | -8.9 |
| **17** | -48.66 | -47.36 | -1.30 | **67** | -44.48 | -38.78 | -5.7 |
| **18** | -48.59 | -36.99 | -11.60 | **68** | -44.48 | -33.38 | -11.1 |
| **19** | -48.54 | -39.84 | -8.70 | **69** | -44.46 | -37.06 | -7.4 |
| **20** | -48.17 | -39.67 | -8.50 | **70** | -44.45 | -44.05 | -0.4 |
| **21** | -47.84 | -43.24 | -4.60 | **71** | -44.45 | -33.65 | -10.8 |
| **22** | -47.39 | -40.49 | -6.90 | **72** | -44.44 | -35.74 | -8.7 |
| **23** | -47.38 | -35.18 | -12.20 | **73** | -44.33 | -41.53 | -2.8 |
| **24** | -47.36 | -45.16 | -2.20 | **74** | -44.33 | -37.33 | -7 |
| **25** | -47.32 | -41.02 | -6.30 | **75** | -44.30 | -39.90 | -4.4 |
| **26** | -47.27 | -44.57 | -2.70 | **76** | -44.25 | -34.95 | -9.3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **27** | -47.23 | -42.63 | -4.60 | **77** | -44.22 | -41.52 | -2.7 |
| **28** | -47.07 | -44.97 | -2.10 | **78** | -44.20 | -41.30 | -2.9 |
| **29** | -46.97 | -43.07 | -3.90 | **79** | -44.16 | -34.56 | -9.6 |
| **30** | -46.91 | -39.11 | -7.80 | **80** | -44.14 | -42.44 | -1.7 |
| **31** | -46.55 | -35.05 | -11.50 | **81** | -44.12 | -42.82 | -1.3 |
| **32** | -46.53 | -42.73 | -3.80 | **82** | -44.10 | -43.40 | -0.7 |
| **33** | -46.49 | -39.59 | -6.90 | **83** | -44.08 | -35.48 | -8.6 |
| **34** | -46.39 | -43.19 | -3.20 | **84** | -44.05 | -37.95 | -6.1 |
| **35** | -46.32 | -42.32 | -4.00 | **85** | -44.04 | -37.24 | -6.8 |
| **36** | -46.23 | -41.03 | -5.20 | **86** | -43.93 | -40.53 | -3.4 |
| **37** | -46.15 | -40.45 | -5.70 | **87** | -43.91 | -36.51 | -7.4 |
| **38** | -46.13 | -43.53 | -2.60 | **88** | -43.91 | -39.21 | -4.7 |
| **39** | -46.04 | -35.94 | -10.10 | **89** | -43.89 | -41.79 | -2.1 |
| **40** | -45.98 | -42.38 | -3.60 | **90** | -43.85 | -37.95 | -5.9 |
| **41** | -45.89 | -37.29 | -8.60 | **91** | -43.82 | -35.62 | -8.2 |
| **42** | -45.87 | -40.47 | -5.40 | **92** | -43.82 | -34.42 | -9.4 |
| **43** | -45.78 | -37.98 | -7.80 | **93** | -43.79 | -38.19 | -5.6 |
| **44** | -45.74 | -42.64 | -3.10 | **94** | -43.77 | -36.27 | -7.5 |
| **45** | -45.72 | -43.22 | -2.50 | **95** | -43.77 | -41.37 | -2.4 |
| **46** | -45.63 | -40.33 | -5.30 | **96** | -43.76 | -40.16 | -3.6 |
| **47** | -45.51 | -40.11 | -5.40 | **97** | -43.76 | -40.56 | -3.2 |
| **48** | -45.48 | -41.08 | -4.40 | **98** | -43.74 | -34.34 | -9.4 |
| **49** | -45.45 | -31.65 | -13.80 | **99** | -43.73 | -38.33 | -5.4 |
| **50** | -45.43 | -40.53 | -4.90 | **100** | -43.69 | -38.19 | -5.5 |

[1]Electrostatic energy
[2]van der Waals energy